

# The Manufacturing Industry and the Big Data Challenges: The case-study of the Global Tech Company

Cezar Honorato<sup>1</sup>, Francisco Cristóvão Lourenço de Melo<sup>2</sup>, Reinaldo Cesar de Morais<sup>3</sup>

<sup>1</sup>Department of Space, Science and Technologies, Technological Institute of Aeronautics, Brazil

<sup>2</sup>Institute of Aeronautics and Space, Technological Institute of Aeronautics, Brazil

<sup>3</sup>Department of Engineering and Science, The University of Salford, UK

Received: 29 Oct 2020;

Received in revised form:

15 Jan 2021;

Accepted: 13 Feb 2021;

Available online: 25 Feb 2021

©2021 The Author(s). Published by AI  
Publication. This is an open access article  
under the CC BY license  
(<https://creativecommons.org/licenses/by/4.0/>).

**Keywords—** *advanced analysis, big data, enterprises, industry and manufacturing*

**Abstract—***Big Data is a crucial component not only to support business decisions, but to gain competitive advantage in an increasingly globalized world, where those who have more information are one step ahead, in this concept companies like technology companies like Facebook, Google and traditional manufacturing companies are making high investments in Big Data and Advanced Analytics with the challenge of translating this huge volume of data. This research effort aims to map the challenges that traditional Manufacturing Industry companies must take advantage of the potential of Big Data and Advanced Analysis to generate real value for the Business through a case study of a global technology company. We found that despite the professionals working on technological development, they understand that the generation of value and the lack of strong evidence of how much Big Data is adding to the current scenario of the Manufacturing Industry was one of the biggest motivations behind this research, which after the case study carried out gave a clear understanding that the manufacturing industry is not prepared to make the most of this new era of Analytics. Another important fact revealed by the surveys is the common concern about the need to invest in improving the competence of the resource in any organization to explore the Big Data environment.*

## I. INTRODUCTION

The Information Chain generated in several steps of the Supply Chain Management from design until after sales have been collected, analyzed and used to support the business decision in several areas of sales, customers behavior, and inventory (Davenport & Dyché, 2013) creating one of the most valuable assets (Knowledge Management) that can define a company competitiveness since the beginning of the modern industry, perhaps even before (Côrte-Real, Oliveira & Ruivo, 2017).

The exploration of the full potential of this scenario is sustained by the technological achievements in data collection, new forms of data organization and analysis that are creating the new environment with large

data of volumes that are also pushing the companies to create capabilities in new areas of knowledge (Malladi, 2013) bringing up the Big Data and Advanced Analytics on the top of all recent investments on technology companies like Facebook, Google, and the traditional Manufacturing companies, with the challenge to translate these huge volume of data (Hu et al., 2014) into strategic decisions as was defined by W. Edwards Deming and Peter Drucker - “You can’t manage what you don’t measure.” (McAfee et al., 2012).

The acronym 3V’s (McAfee et al., 2012) was one of the first ways to describe the Big Data characteristics defining them as volume – The quantity of data extracted, stored and analyzed, Velocity – Real-Time data, data

extracted and transported as is generated and Variety – text, messages, images, sensors, videos, etc. Other characteristics have been added by different theories and studies lately, like Veracity – data reliability and data sources credibility (Abbasi et al., 2016), Value – to measure value generated by the Big Data to support the Business Decision – and Complexity – the level of interconnections and interdependencies between the several data structures (Kaisler et al., 2013). To be able to generate value and enhance competitiveness it is necessary to manage the Business Data Analytics managing its characteristics properly (Wamba et al. 2015).

Nevertheless, the benefits of Big Data implementation have been researched and studied heavily last ten Years, even before. A simple research in Google Scholar using the keyword “Big Data” during the Year of 2017 returns more than 65.700 results. The companies are trying to enhance their organization's skills in the Data Scientist area in order to be able this handle the new volumes of data, what can be much more challenge than the traditional quantitative analysis that has been used historically to support the business decisions. New competencies in the Data Science area are required (Davenport & Dyché, 2013) to understand the value that can be generated by the data exploration, run the predictive and prescriptive analysis, “tell a story with data” and explore deeper than what the first layer of data is showing and take advantage of the potential value realization (Agrahri et al., 2017) that the Big Data Era is promising.

The Return on Investment between investments in traditional Data Warehouse and a Big Data environment in large companies was published by Davenport & Dyché (2013) showing a clear advantage of investments in Big Data with return on cash flow and time to value, but the lack of real and rigorous proof of return of investments generated by Big Data investments brings us the conclusion it is a long-term strategy taken by the company's Senior Management.

In this way, we have the research question that greatly impacts the manufacturing enterprises: How is the Manufacturing industry preparing itself to take advantage of Big Data environment to generate value? The aim of this research effort is to map the challenges that traditional Manufacturing Industry companies have to take advantage of the Big Data and Advanced Analytics potential generating real value to the Business.

All of these aspects are the reason to conduct this study and look for how to understand the Manufacturing Industry challenges to generate value that is advertised by the Big Data community.

This study will be valuable to support the Manufactory Industry journey in the Big Data area providing the experience and best practices in the company preparation and enhance the knowledge that is important to reduce the time to collect the value generated by such investment, in technology and knowledge.

## II. LITERATURE REVIEW

In the various researched sources, we will make a critical assessment of the different approaches, theories and findings reported by researchers and technical publications flowing through different areas related to Big Data and Big Data Analytics. First, we will assess the definitions of Big Data and how different authors and companies are driving this impulse, secondly, we will do an analysis of publications and studies with the objective of understanding and defining how companies are approaching Big Data Environment. Data to support your investments in this area, thirdly we will analyze the definitions of value generated by investments in Big Data, evaluating different models proposed to measure the effects of Big Data on the company's performance and on the experiences reported by companies through case studies of the same. implementation in different business sectors.

The volume of such data generated and captured by different areas has been increasing exponentially (Agarwal & Dhar, 2014) due to several reasons but mostly driven by the technical achievements in telecommunication and mobility with the advent of devices like RFID, IoT sensors, and mobile phones, as examples, supported also by the strong costs reduction in data storage with the cloud option offered by different companies.

Estimations have been done by Cisco predicting that the number of electronic connections is surpassing the number of people by 2020 (Evans, 2011) as the same said by Gartner (Gartner, 2014) that we will reach 26 Billion of devices in this planet generating 40 ZettaBytes (1 ZettaByte stands for 1 Trillion GigaBytes) (Gantz & Reinsel, 2012) growing up to 163 ZettaBytes by 2025, ten times more the estimated data generated in 2016 (Reinsel, 2017) revealing an enormous potential of exploration and revenue generation. However, it won't be easy to explore and take the advantages of this business opportunity looking due to the variety of how this data is generated currently, as it has been the most common as structured data way in text/numeric format or going to a totally new variety as unstructured like audio, video, images, etc. (Silvarah et al., 2017).

Nevertheless, what is Big Data after all? What is the difference if we compare the current data landscape against the traditional definition of data? We don't have a certain origin of the definition and the concept of Big Data, but (Diebold et al., 2012) argues that this definition probably was coined during a conversation at Silicon Graphics at the 90's and have been promoted by companies like IBM and other leaders in this area since there. (Laney, 2001) was one of the firsts to use the acronym 3V's to describe the main characteristics of Big Data: Volume/ data size in terms of storage and quantity of information (order of Zettabytes ZB), Variety/ structure (texts), semi-structured and unstructured data (audio, video, image, etc.), and Velocity/ real-time data streaming, data collected as is generated. More characteristics have been added later by other scholars, like Value/ information value extracted from the data and Veracity/ Data reliability.

Another ways to reference the data understanding layer when we talk about the tools and technologies used to explore the information relying on the BD layer is Business Intelligence (BI) or Business Analytics (BA) that has been nominated as part of the Top priorities for CIO's in Gartner annual survey in three of the five years between 2007-2011 (Seddon et al., 2012) creating the believe that it will make a remarkable contribution to leverage firms performance in the coming years. As defined by Wieder & Ossimitz, (2015) BI is the analytical layer used to interpret the raw data through software products used to manage the databases until to generate the analysis and create the visualization dashboards used to investigate and support the business decision.

Despite of the different definitions of how to describe Big Data (BD) using its primary characteristics based on data volumes or based on the analytics layer (BI, BA or BDA) it is a common sense that this new scenario is bringing new opportunities and also challenges to be properly explored by the companies to generate value. It is heavily promoted by technologies companies that are developing the framework and several positive business cases especially using the social media and public information, but it is still an area to be explored and validated by the manufacturing industry.

That Big Data is a reality and has an enormous potential to leverage business competitiveness in different areas that have been promoted, researched and discussed last decade is a fact. It is also a fact that it is bringing the advent of new technologies, requiring new competencies and pushing the company's mindset on how to use properly the data owned by the company combined with public information. In order to be able to handle this environment, it is necessary to have the right framework.

Malladi, 2013) used the TOE (Technology – Organizational – Environment) approach to evaluate the framework challenges faced by the companies to implement the tools and systems used to support the Big Data Analytics evaluating the benefits, organization readiness, and internal technology competencies.

According to Akerkar (2013) and Zicari (2014) this framework can be defined based on the challenges through the data life cycle starting from the data characteristics, the process in how to handle the data and the data governance in the aspects of privacy, security, and ethics. However, Silvarah et al. (2017) define the Big Data framework using the a different approach, also based on the data lifecycle but naming five different steps instead three: Data Acquisition and Warehousing, Data Mining and Cleansing, Data Aggregation and Integration, Data Analysis and Modelling and Data Interpretation once (Bizer et al., 2012) named six steps: capturing, storage, searching, sharing, analysis, and visualization. Nevertheless, how we identify and name each phase since the beginning, each of these steps has its challenges once that the data grow in volume and diversity coming from several sources (Paris, Donnal, & Leeb, 2014).

The first challenge that is faced by the companies that are looking for to explore this new world is the technology necessary to handle this data. It not just a matter of acquiring data but also store and secure the information what could demand a high level of investment due to the cost of infrastructure (Wang & Wiebe, 2016). The decision to move to a cloud storage area offered today by several companies is an alternative to get down costs for this infrastructure and the company lack of expertise to administrate this environment. However, this option will bring up another concern about data privacy, performance, and security (Muriiti & Kotzé, 2013).

There is no silver bullet in this choice, the decision must be taken by the company based on its investment capacity, knowledge base in this area and which risks being assumed. This is all about the first part of this environment that has been defined by (Malladi, 2013) as a crucial knowledge for the company ambition. On top of that, (Popovič et al., 2012) emphasize that it is also important to establish the analytical capabilities of the organization that will process and execute the analysis.

Even more challenges have been advised and described by different researchers and published white papers based on several experts in this area, like is well described in the study published by Heyns (2015) on how to prepare the organization structure, data security, creating the high insight and the necessary talent to convert data into information, the right information used in

the right way. This survey indicates that 56% of the companies have concerns to adapt their organizational structure and 50% to find the suitable skills necessary, the required competence. What to do with the huge volume of data and variety of information collected and stored in these big capacity data systems in the owned company IT infrastructure or in the cloud?

This question brings us to the second layer of the Big Data Framework: Analysis. This is recognized as a valuable and unreplaceable capability that drives the organization towards the competitive advantage unlocked by the Big Data Analytics (Wamba et al, 2017). One of the premises to move towards the Big Data Analytics is the possibility to enhance business decisions moving from the reactive traditional descriptive ways of working ("What happened?") to take advantage of its different analytics possibilities and becoming more predictive and prescriptive what can turn into a quite complex process that requires the right competencies inside the organization or using companies that provide such kind of service (Assunção et al., 2015).

In this case, the challenge is how to create the domain knowledge that is necessary to look for the correct answers when exploring the process data, (Waller & Fawcett, 2013) described this as the difficult to teach the related area knowledge that is created with the experience as we can teach the technical tools usage, like statistical programming languages like Python or R what can be one of the reasons for implementation failures faced by the companies. In the (Genpact & Fortune, 2017) survey only 38% of the company participants said that they are preparing their employees with training opportunities or reskilling options in these new technologies. Data Analysis is an area that has been evolving exponentially last 10 Years

All of the aspects described here have been discussed and researched by several scholars lately with the common indication that the implementation of Big Data Analytics is not a riskless journey that demands an enterprise heavily planning and preparation (Côrte-Real, Ruivo & Oliveira, 2017). However, the success only can be achieved with the combination of the company knowledge base together with the combination of the new modern technologies and approaches like the ELTA (Extract, Load, Transform and Analyze) and the well tested and proved existent techniques (Marín-Ortega et al., 2014).

### III. METHOD

#### A. Research Philosophy

The nature of this study is supported by the philosophy of the Critical realism combined with the abduction approach

once that the main objective is to understand and explain observable events and information collected during the interviews. A Qualitative research was the method decided to be applied in this study to support the research question investigation and conclusion. Qualitative research is based on different ways to collect non-numerical data from small samples once that Quantitative method is based on numerical data normally collected from larger sources of information (Saunders, Lewis & Thornhill, 2009).

Case Study is one of the genres of a Qualitative research and creates the right environment to allow a deep investigation of a particular event (Saldana, 2011). This decision was made based on the characteristics of this study when we compare with the four reasons described by (Yin, 2003) that support a Case Study design consideration: a "how" research question, the participants behavior cannot be manipulated, the contextual conditions are important to the phenomenon under study the unclear overlay between the phenomenon and the context of the study.

#### B. Research Strategy

A Case Study can be defined as a mixed approach in terms of to be inductive or deductive because involves real events investigations in order to understand a specific contemporary phenomenon (Datt, 2016) therefore the following strategies were adopted.

Our research strategy consists of three types of procedures: (i) Documentary: The company data structure before the Big Data implementation was described and the main constraints that drive the decision of the change delineated. The new data architecture and the achievements expected with this implementation compared with the real results. The observation will be also important to describe the challenges and how the organization was prepared to face them. (ii) Internal Survey: The Individual opinion from experimented professional from different areas of expertise involved directly and indirectly in the Big Data implementation project was collected in form of a questionnaire to understand how the organization knowledge and perception of such change are aligned and coordinated between heterogeneous background and responsibility. A questionnaire based on closed questions was used to limit the amplitude of possible answers and as the most effective way to gather concrete information from the participants (Ranjit, 2011) and (iii) External Survey: An external survey was used to capture how other companies from different sectors are seeing the Big Data revolution, their intention to invest in this area or not, what how they perceive the potential benefits and challenges correlated to this area.



### C. Research Design

The research was designed in a way to make possibly understand the company approach to its Big Data implementation project, collect the achievements targeted with this implementation and the perception of the main resources involved in this area about the challenges faced by the organization. On top of that, an external Survey was released to also capture the opinion of professionals involved in similar projects at another company.

The following details of each step of the Case Study process as was defined by (Yin, 2009) was applied for this study realization. The Plan phase was mainly composed by the search of the relevant literature in this area to support and justify the research motivation and the understanding of the company investments in this area to be the right object of study. The Design phase was focused on to decide how to collect the data necessary to support the conclusions to answer the research question, how to understand the company environment and the creation of the survey.

The prepare phase was focused on the test the surveys and identify the channels to be used to share the surveys and collect the data. The data collection based in the surveys was done using two different channels, one internal and another one external exploring the researcher professional networks, the organization implementation project was mapped to make possible to compare the environment before the implementation and correlate this initiative with the data collected through the surveys. After that, data analysis was done looking for to identify the information collected and how to conclude if the research question was properly answered and what are the study conclusion. The result was shared with the organization and also compiled to be sent to all participants of the external survey that register the interest to have it.

To answer our research question, we have some objectives to be fulfilled, namely:

1. To understand the investments to capture, inject, store and explore the data
2. To understand the new competencies required by this new scenario
3. To understand how the Manufacturing Industry is preparing its organization to embrace this change
4. To understand the differences between Big Data exploration experience at technology companies and Manufacturing Industry
5. To define how to measure the Business Value of Big Data in the Manufacturing Industry.

In this way, the research questionnaire was elaborated based on eleven questions in order to guide the understanding of the proposed goals.

Table 1: Survey questions

Nr	Survey Type	Question	Objective	Goals
1	Internal	What is your area of work currently?	Identify the participant area of work	
1	External	What is your company primary area of actuation?	Capture the participants companies' business area	
2	Internal/External	What is your understanding about Big Data and Advanced Analytics like Machine Learning and Artificial Intelligence?	Identify the participant level of knowledge about Big Data and Advanced Analytics	
3	Internal	Have you been involved in any project connected to Big Data or Advanced Analytics recently?	Map the participant experience in Big Data and Advance Analytics projects	
4	Internal/External	How much do you agree with the statement: "Information is the Oil of new era - (Gartner)"	Understand the participant's opinion about the importance of Big Data in the actual scenario	3
5	Internal/External	How much do you agree with the statement: "Big Data is applicable to any kind of Business"	Understand the participants level of knowledge about Big Data and Advanced Analytics	2,3
6	Internal/External	How much do you agree with the statement: "The Manufacturing Industry is prepared to explore Big Data environment"	Understand how the companies perceive themselves prepared to explore Big Data environment	2,3
7	Internal/External	How much do you agree with the statement: "Big Data value generation is strongly tested and proved in the Manufacturing Industry"	Understand the participant's perceptions of Big Data successful implementation knowledge	2,3
8	Internal/External	How much do you agree with the statement: "The Manufacturing Industry has the appropriate Analytics Maturity Level to explore the Big Data Environment"	Understand the Participant opinion about the Manufacturing Industry Analytics stage	2,3,4

In order to answer the research question already presented in the introduction: How is the Manufacturing Industry preparing itself to take advantage of Big Data environment to generate value? Based on this research question, we elaborate the hypothesis: Is the Manufacturing Industry prepared to implement and generate value from Big Data and Advanced Analytics technologies?

To collect the data necessary to validate the hypothesis and evaluate the objectives connected to this question we follow the procedure: Firstly, all the information about the organization object of this study was collected to understand the original environment, the project ambition, the implementation plan, the challenges faced, and the results achieved. Secondly, an internal survey was released to collect the opinion from worldwide professionals involved directly and indirectly in this project regarding the theme and the organization approach. Thirdly, an external survey was released to capture the

opinion from different sectors of manufacturing and correlate them with the internal survey data and the project information.

#### D. Case-Study: The Company Description

The company object of this study is a high technology equipment manufacturer with production facilities in several places around the world using a well-established IT network to collect production information from its own sites and from third part suppliers producing on behalf of this company based on supplied design and test environment. These data providers delivery thousands of records every minute twenty-four hours per day to support local and global performance reporting and analysis (Figure 1).

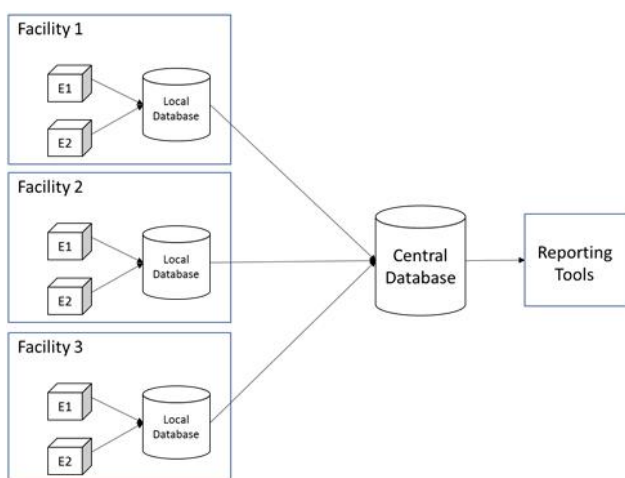


Fig. 1: Company Production Data IT structure

The data flow from the production and test equipment to the local databases that create hourly batches to be transmitted towards a security channel with a latency between two and four hours depending on production volumes, which type of data and the facility location. All the information retrieved from the production facilities is used to monitor the Product Quality, Production performance, support Root Cause Analysis when it is necessary and secure product traceability linking the products with the results and components used during the manufacturing process. This architecture is based on traditional SQL Databases technology and demand a high performance and capacity hardware to secure that the users will be able to explore such data without slowness and data quality issues. Therefore, data reliability and quality are crucial aspects to secure the investigation and decision required by the business

The Data Architecture described before proves to be efficient to provide to the company the right environment that was necessary in order to conduct the manual diagnostic and investigative daily work done by

hundreds of Quality and Product Engineers in all of these facilities and the Product Designers interested to follow up the product performance comparing volumes results against design expectation. However, with the explosion of Big Data and Advanced Analytics tools providing the analysis capabilities and the hardware technology evolution providing the computing capacity to evaluate thousands of measurements per second a new scenario started to be achievable by the manufacturing industry that is under a fierce competing in this area looking for new ways of reducing costs and increase its competitiveness (Figure 2).

The Big Data implementation project had the main following objectives:

1. Reduce the data latency from 2-4 hours to not more than 30 minutes
2. Provide the right environment to run advanced analytics tool

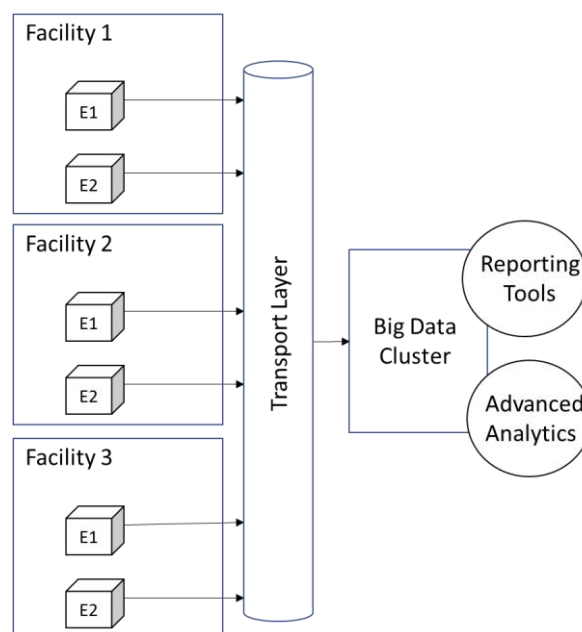


Fig.2: Big Data infrastructure design

This new infrastructure would create the right environment to support the company ambitions to reduce production costs worldwide moving from the traditional descriptive and diagnostic ways of work to predictive and prescriptive analysis detecting anomalies and not expected behavior as early as possible to analyze and react reducing losses with bad quality and production shortage risks.

Due to the confidentiality and the objective of having more objectivity in the answers, the participant's name data will not be requested and the company, all the participants contribute to this study voluntarily.

#### IV. FINDINGS

The first part of our findings was based on the observation and collection of information regarding the Big Data implementation project driven at the company under study. We follow the project implementation since the beginning what help us to understand how the company address our research questions objectives by topic.

About the first objective proposed: “To understand the investments to capture, inject, store and explore the data.” we were not able to understand this point once that the company under evaluation already have the necessary infrastructure available by other areas to be explored and was not necessary to make specific investments in infrastructure. Nevertheless, it is an advantage if the company is able to share infrastructure between different areas.

In the second objective: “To understand the new competencies required by this new scenario”. The company defined a couple of specific Use Cases to explore the data and identify what kind of advanced analytics and machine learning algorithm could be used to run predictive analysis and generated the target value. This work demanded a new cross organization team composed by resources with knowledge and data access, the product domain knowledge and data scientists with the required knowledge to apply the tools and technologies available in the market to execute this work.

About the third objective “To understand how the Manufacturing Industry is preparing its organization to embrace this change”, another aspect of this implementation was the idea promotion through intensive discussions involving different areas of the company in order to align the understanding of the organization around the theme, get the right attention, secure the necessary engagement from the right stakeholders and understand the Analytics company strategy to secure the project alignment.

As regards the proposed fourth objective “To understand the differences between Big Data exploration experience at technology companies and Manufacturing Industry”, we could notice that one of the major differences between the Manufacturing Industry and the technology companies is the data characteristics. The technology companies explore unstructured data using several different machine learning algorithms with high focus on customer understand (external) once that the manufacturing data is more composed by structure data with attention in process performance and production quality (internal), however the usage of unstructured data like image, IoT devices and binary large objects is also increasing in this environment.

And as the fifth and final proposed objective “To define how to measure the Business Value of Big Data in the Manufacturing Industry”, in this specific case, the expected value generation was connected to detect test results anomalies during the manufacturing process and take the corrective actions to adjust the performance to the desired level without reduce the production throughput and affect the product quality. However, we could not quantify and validate this achievement because the project was not finished before we conclude this study. The theory was tested and proved using historical data extraction where the models were created and tested ratifying the possibility to use this approach, but it was not possible to testify this structure up and running.

The second part of the findings were based on our Internal and external Survey answers compilation. The Internal survey was sent to 95 professionals that were directly or indirectly involved in the Big Data implementation project, another area of Analytics and Product Quality reports consumers. The response rate was 55% which means that we got 62 answers. The external survey it was sent for 95 professionals and was received 56 answers, the response rate it was 59%.

About the question 1 of the questionnaire: What is your area of work? (Internal) / Company Business area (External) the results are shows in the chart1. The vertical bars show the number of participants per work area (scale to the left of the graph) and the lines the respective accumulated percentage (scale to the right of the graph), for example, in the internal participants of the company, in the test development area we had the participation of 14 professionals representing approximately 20% of the total, the results are shows in the figure 3.

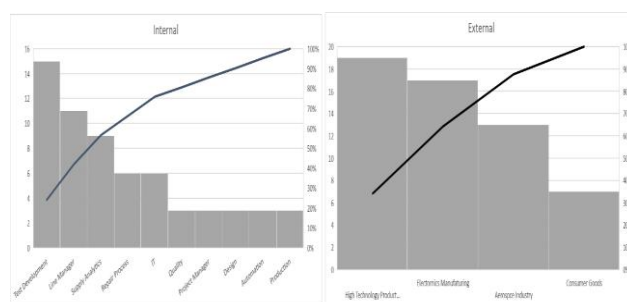


Fig.3: Participants area of work

As regards the question 2 of the questionnaire: What is your understanding about Big Data and Advanced Analytics like Machine Learning and Artificial Intelligence? the results are shows in the figure 4.

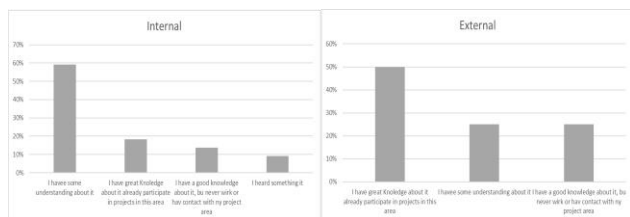


Fig.4: Participants level of understanding about Big Data Technologies

On question 3 of the questionnaire: Have you been involved in any project connected to Big Data or Advanced Analytics recently? the results are shown in the figure 5 (Internal only).

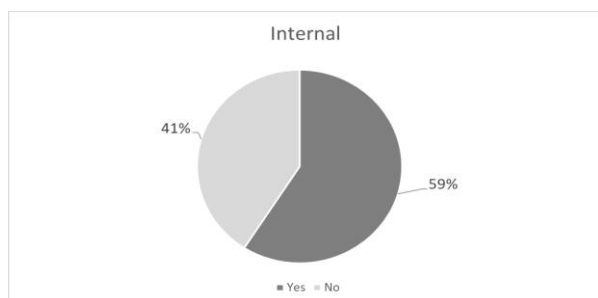


Fig.5: Big Data Project participation

The three first questions were designed to create a good understanding of the participants' involvements in our research area and their level of knowledge of Big Data technologies. We can conclude from the answer's compilation that most of the participants have a good understanding and have been involved in different projects in this area.

About the question 4 of the questionnaire: How much do you agree with the statement: "Information is the Oil of new era - (Gartner)" the results are shown in the figure 6.

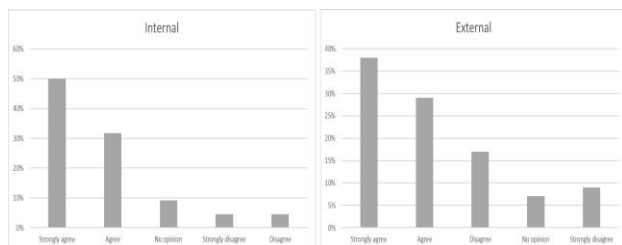


Fig.6: "Information is the Oil of new era (Gartner)"

In the question 5 of the questionnaire: How much do you agree with the statement: "Big Data is applicable to any kind of Business" the results are shown in the figure 7.

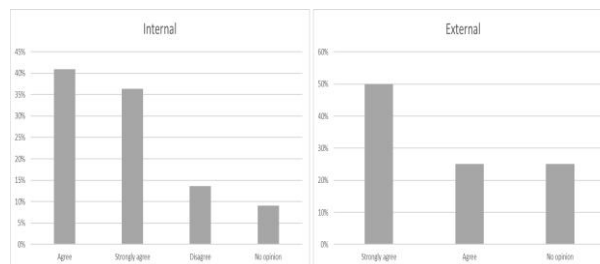


Fig.7: "Big Data is applicable to any kind of Business"

As regards the question 6 of the questionnaire: How much do you agree with the statement: "The Manufacturing Industry is prepared to explore Big Data environment" the results are shown in the figure 8.

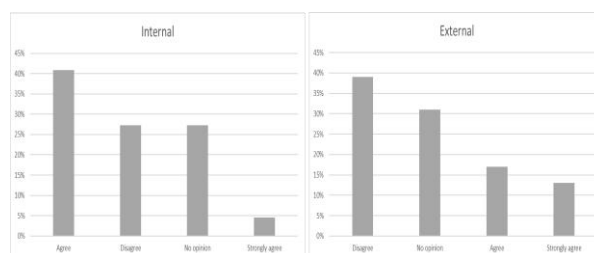


Fig.8: "The Manufacturing Industry is prepared to explore Big Data environment"

About the question 7 of the questionnaire: How much do you agree with the statement: "Big Data value generation is strongly tested and proved in the Manufacturing Industry" the results are shown in the figure 9.

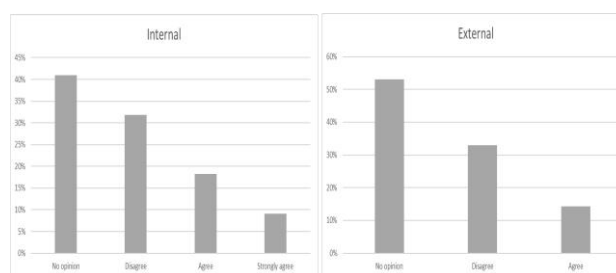


Fig.9: "Big Data value generation is strongly tested and proved in the Manufacturing Industry"

In the question 8 of the questionnaire: How much do you agree with the statement: "The Manufacturing Industry has the appropriate Analytics Maturity Level to explore the Big Data Environment" the results are shown in the figure 10.



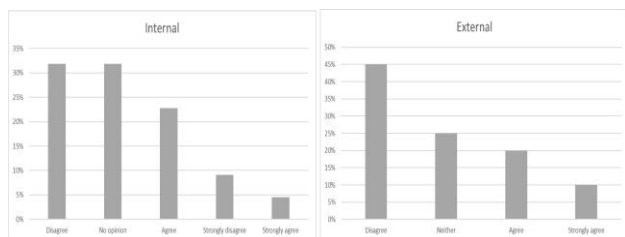


Fig.10: Manufacturing Industry maturity to explore Big Data

The questions four to eight were designed to capture the participants opinion of how Big Data and Advanced Analytics is perceived as an important asset to generate value, how it could be applied to the Manufacturing Industry environment, what is the Manufacturing Industry stage in terms of knowledge and how the Manufacturing Industry perceives their maturity to use predictive and prescriptive automatic analysis instead the ordinary manual diagnostics ways of work. It is important to notice that there is a strong alignment between the internal survey done at the company under study and the external survey despite the low number of samples collected.

About the question 9 of the questionnaire: Which area do you believe that Big Data can be most applicable to the Manufacturing Process? the results are shows in the figure 11.

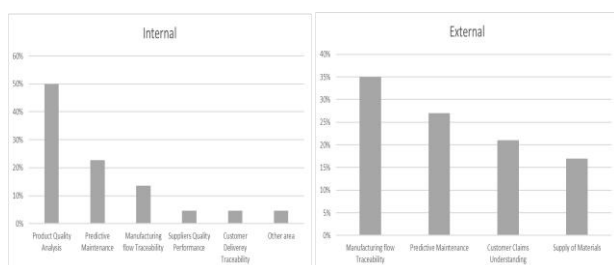


Fig.11: Manufacturing Industry applicable areas

As regards the question 10 of the questionnaire - Which layer do you believe is most challenge to achieve the promised Big Data results? (Internal only) the results are shows in the figure 12.

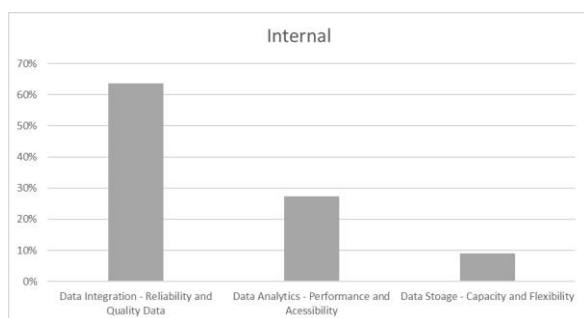


Fig.12: Layers challenges

About the question 11 of the questionnaire: Which area do you believe that the companies are facing most critical challenges in the Big Data implementation? the results are shows in the figure 13.

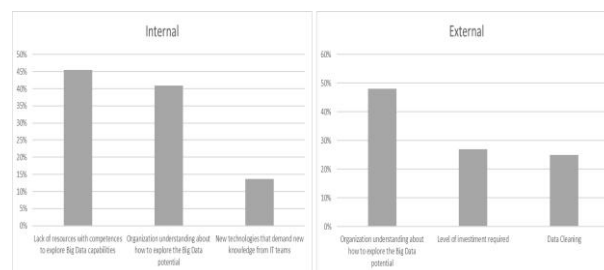


Fig.13: Manufacturing Industry Challenges

In the question 12 of the questionnaire: In your opinion, what is the most important area to receive major investments from the Manufacturing Industry to explore the full potential of Big Data and Advanced Analytics? the results are shows in the figure 14.

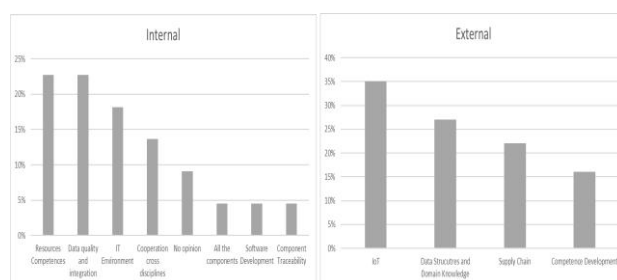


Fig.14: Major areas of investment

Questions nine to ten were designed to capture the participants vision about areas of future investments, Big Data and Advanced Analytics applicability in the Manufacturing Industry and major areas of challenge. It is strongly noticed that the right competences to take advantage of the data exploration is an important area to be seriously addressed if the company wants to be successful in this area.

## V. CONCLUSION

As this research effort is based on a case study of a Big Data and Advanced Analytics implementations following an implementation project in a chosen company combined with the application of internal and external surveys to capture information and data that could be used to validate our defined hypothesis: Is the Manufacturing Industry prepared to implement and generate value from Big Data and Advanced Analytics technologies?

It was mainly motivated by the high attention from different sectors of the Manufacturing Industry that is driving high level of investments (Bhadani & Jothimani, 2016) recently and in the coming years and the lack of strong evidence about the value generation provided by Big Data Analytics in this business segmentation.

The decision to follow a project implementation in a Global Company with a recognized and robust IT architecture supporting the ordinary ways of work with data analysis was a good approach because it was possible to understand the advantages and disadvantages of to move from a tested and well proved environment to new ways of work instead start a full implementation since the beginning without a strong experience and mature organization working with a different environment. Another good point of this approach was to have the possibility to contact several professionals from different sectors and background to understand this challenge from different perspectives. From another hand, the time was a crucial limitation to support this research to reach conclusions in areas like investments and value generation.

The idea of the external survey was explored to understand how the company object of the case study professional's vision and experience regarding the theme is aligned with professionals with the same or higher level of experience from another manufacturing business areas to validate their understanding, challenges an expectation.

The survey's results show a strong alignment when the answers are compared side by side and statistically tested and could be a valuable tool to ratify the hypothesis conclusion if a more statistical relevant participant were achieved by both surveys, internal and external. The weakness point of the external research is the lack of possibility to certify the participants level of experience and knowledge regarding the theme.

Nevertheless, we strong believe that the facts and data collected during this work despite the challenges to complete some of the planned investigation and data collection provide a very good and reliable piece of work that is extremely relevant to contribute to the community base of knowledge and recommend directions for another company's investments in this area.

One of the aspects of this research to understand the needed investments in Big Data implementation compared to traditional warehouse architecture was mainly motivate by the divergent opinions and results from different studies and was already promoted by (Davenport & Dyche, 2013) that the Big Data environment investments has a better return of investments when compared to traditional warehouse architecture and the intense discussion around the "IT Productive paradox" with several

other studies showing that not always this kind of investment is able to bring value to the organization (Anand, Fosso, & Sharma, 2013; Frisk, Lindgren & Mathiassen, 2014).

To understand how the Manufacturing Industry is preparing its organization to embrace this change it was collected good point by the surveys answers was the company understanding about the potential of Big Data as "The oil of new era" (Yi, Liu, Liu, & Jin, 2014) and its applicability to any kind of process and different areas of the Manufacturing process as it was clearly appointed with the alignment of all participants answers confirming previous theories and different business investments in this area (Bhadani & Jothimani, 2016). The results also show us that this has been a topic of discussion and an area with several projects in different companies due the level of understanding about the theme demonstrated by the participants internally and externally.

To understand the differences between Big Data exploration experience at technology companies and Manufacturing Industry, the main point identified in this area was which part of the process the Manufacturing Industry could get the best benefits when applying Big Data Analytics when we compare with retail and technology companies where customer consumption trends and behavior are the mains addressed areas (McAfee et al., 2012) together with inventory management (Chen & Zhang, 2014) once that in the Manufacturing Industry the major concerns are on the internal process optimization and costs reduction.

However, we cannot simply assume it as the only areas to be addressed once that our survey participants are mainly from product quality and manufacturing process and not from Customer Units area or Marketing where probably we could get a different result. We got some inputs from our surveys that the Supply Chain Information management (ISM) (Agrahri et al., 2017) is also an area relevant to secure the Manufacturing Industry competitiveness and deserves to be in the agenda for investments on Suppliers Quality, Products and Customer delivery traceability.

We can conclude that even though we were not able to prove and justify all objectives in our research questions the overall understanding and some theories evaluation were successfully addressed in this work bringing a valuable contribution applied in the real technology company to the community knowledge base.

## ACKNOWLEDGEMENTS

We appreciate the recommendations of the evaluators who actively contributed to the reflection, expansion and

development of the ideas of this work. This research effort was supported by National Counsel of Technological and Scientific Development (CNPQ) and Coordination for the Improvement of Higher Education Personnel (CAPES). We greatly thank for supporting this research.

## REFERENCES

- [1] Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the association for information systems*, 17(2), 3.
- [2] Agarwal, R., & Dhar, V. (2014). Big data, data science, and analytics: The opportunity and challenge for IS research.
- [3] Agrahri, H., Ahmed, F., Verma, V. K., & Purohit, J. K. (2017). Benefits of implement big data driven supply chain management: an ISM based model. *International Journal of Engineering Science*, 7(5), 11426-11431.
- [4] Akerkar, R. (2013). Advanced data analytics for business. *Big data computing*, 377(9).
- [5] Anand, A., Fosso Wamba, S., & Sharma, R. (2013). The effects of firm IT capabilities on firm performance: the mediating effects of process improvement.
- [6] Assunção, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A., & Buyya, R. (2015). Big Data computing and clouds: Trends and future directions. *Journal of Parallel and Distributed Computing*, 79, 3-15.
- [7] Bhadani, A. K., & Jothimani, D. (2016). Big data: challenges, opportunities, and realities. In *Effective big data management and opportunities for implementation* (pp. 1-24). IGI Global.
- [8] Bizer, C., Boncz, P., Brodie, M. L., & Erling, O. (2012). The meaningful use of big data: four perspectives--four challenges. *ACM Sigmod Record*, 40(4), 56-60.
- [9] Chen, C. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information sciences*, 275, 314-347.
- [10] Côte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of Business Research*, 70, 379-390.
- [11] Datt, S. (2016). Defining research strategy in a research paper on business studies. *Project Guru*.
- [12] Davenport, T. H., & Dyché, J. (2013). Big data in big companies. *International Institute for Analytics*, 3, 1-31.
- [13] Diebold, F. X., Cheng, X., Diebold, S., Foster, D., Halperin, M., Lohr, S., ... & Shin, M. (2012). A Personal Perspective on the Origin (s) and Development of "Big Data": The Phenomenon, the Term, and the Discipline\*.
- [14] Evans, D. (2011). The internet of things: How the next evolution of the internet is changing everything. *CISCO white paper*, 1(2011), 1-11.
- [15] Heyns, H. (2015). Becoming an analytics-driven organisation to create value. *Ernst & Young, United Kingdom, Technical document*.
- [16] Frisk, J. E., Lindgren, R., & Mathiassen, L. (2014). Design matters for decision makers: Discovering IT investment alternatives. *European Journal of Information Systems*, 23(4), 442-461.
- [17] Gantz, J., & Reinsel, D. (2012). The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. *IDC iView: IDC Analyze the future*, 2007(2012), 1-16.
- [18] Gartner, W. W. W. (2014). Gartner says the Internet of Things will transform the data center. Retrieved from <http://www.gartner.com/newsroom/id/2684616>
- [19] Genpact, G., Fortune, G. (2017) "Is Your Business AI-Ready?" Retrieved from <https://www.genpact.com/downloadable-content/insight/is-your-business-ai-ready.pdf>
- [20] Hu, H., Wen, Y., Chua, T. S., & Li, X. (2014). Toward scalable systems for big data analytics: A technology tutorial. *IEEE access*, 2, 652-687.
- [21] Kaisler, S., Armour, F., Espinosa, J. A., & Money, W. (2013, January). Big data: Issues and challenges moving forward. In 2013 46th Hawaii international conference on system sciences (pp. 995-1004). IEEE.
- [22] Laney, D. (2001). 3D Data management: Controlling data volume, velocity and variety. Meta Group. *Lakshen, Guma Abdulkhader*, 1-4.
- [23] Malladi, S. (2013). Adoption of business intelligence & analytics in organizations—an empirical study of antecedents.
- [24] Marín-Ortega, P. M., Dmitriyev, V., Abilov, M., & Gómez, J. M. (2014). ELTA: New approach in designing business intelligence solutions in era of big data. *Procedia technology*, 16, 667-674.
- [25] McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. *Harvard business review*, 90(10), 60-68.
- [26] Muriithi, G. M., & Kotzé, J. E. (2013, October). A conceptual framework for delivering cost effective business intelligence solutions as a service. In *Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference* (pp. 96-100).
- [27] Paris, J., Donnal, J. S., & Leeb, S. B. (2014). NilmDB: The non-intrusive load monitor database. *IEEE Transactions on Smart Grid*, 5(5), 2459-2467.
- [28] Popović, A., Hackney, R., Coelho, P. S., & Jaklič, J. (2012). Towards business intelligence systems success: Effects of maturity and culture on analytical decision making. *Decision Support Systems*, 54(1), 729-739.
- [29] Ranjit, K. (2011). RESEARCH METHODOLOGY a step-by-step guide for beginners.
- [30] Reinsel, D., Gantz, J., & Rydning, J. (2017). Data age 2025: The evolution of data to life-critical. *Don't Focus on Big Data*, 2.
- [31] Saldana, J. (2011). *Fundamentals of qualitative research*. OUP USA.
- [32] Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research methods for business students*. Pearson education.
- [33] Seddon, P. B., Constantinidis, D., & Dod, H. (2012). How does business analytics contribute to business value?.
- [34] Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and

- analytical methods. *Journal of Business Research*, 70, 263-286.
- [35] Zicari, R. V. (2014). Big data: Challenges and opportunities. *Big data computing*, 564, 103.
- [36] Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management.
- [37] Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, 234-246.
- [38] Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70, 356-365.
- [39] Wang, Y., & Wiebe, V. J. (2016). Big Data Analytics on the characteristic equilibrium of collective opinions in social networks. In *Big Data: Concepts, Methodologies, Tools, and Applications* (pp. 1403-1420). IGI Global.
- [40] Wieder, B., & Ossimitz, M. L. (2015). The impact of Business Intelligence on the quality of decision making—a mediation model. *Procedia Computer Science*, 64, 1163-1171.
- [41] Yin, R. K. (2003). Case study research: design and methods (ed.). *Applied social research methods series*, 5.
- [42] Yin, R. K. (2009). Case study research: Design and methods fourth edition. *Los Angeles and London: SAGE*.
- [43] Yi, X., Liu, F., Liu, J., & Jin, H. (2014). Building a network highway for big data: architecture and challenges. *Ieee Network*, 28(4), 5-13.